2

Analysis II

Mohammed Fulano

Analysis II

Introduction to function spaces

Mohammed Fulano

B.Sc. in Applied Mathematics

Algeria

First version : 10/09/2018

The book still being updated, any mistakes occurring here are mine.

Preface

These notes are considered as the part that should follow to the previous book with the title *Analysis I: infinitesimal calculus*. They are an introduction to the methods of functional analysis used to solve problems about differential and integral equations. Functional analysis is seen as linear algebra for infinite dimensional spaces. Especially it is the study of certain linear spaces, the *function spaces* and their use in the solution of the ***Cauchy problem*** associated to those equations.

The search for existence and unicity of a solution for boundary problems, and all the problems turning around it, both applied and theoretical, are quite old and taught in all the universities worldwide. These problems should be solved as they occur in many practical issues, in engineering, physics, economics, biology,…

So to get to these important studies, mathematicians begin first by simple things as finite dimensional linear spaces, and as usual when they enter in the infinite case, they discover many dramatic changes.

To move from simple calculus toward all this complicated machinery, students need three essential tools in undergraduate curriculum:

−Infinitesimal calculus (i.e. differential and integral calculus).

−Linear algebra.

−General topology(point set topology).

My aim is especially to give the reader the logic of the transition from the calculus to functional analysis and the difficult points he may met through this itinerary, so the chart of my notes to reach this aim will be as follows:

Chapter 1:

In this first chapter we begin with the theory of linear spaces studied in Linear Algebra and its ideas. This theory with its linear aspects is very important in many branches of mathematics for its simplicity compared to nonlinear aspects. For instance linear spaces are the spaces with which we will work all the time. Even if the space is not metric, its linearity will help us a lot. Here we see only finite dimensional spaces beginning by the very natural space $\mathbb{R}^n$. We will review vectors, vector spaces, linear dependence, basis ...

Chapter 2

We study norms and their generalization to metrics, the idea is that we want to generalize the length of a vector, like the length of a segment in the three known spaces $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3$, to the general space $\mathbb{R}^n$. We first begin by taking a short look at general neighborhoods studied in general topology, then we move to metrics then to norms. Norms as special cases of metrics, are the very natural to work with, and of course the first natural norms in $\mathbb{R}^n$ are presented. Later we learn that the norms in infinite dimensional function spaces are very tricky.

Chapter 3:

To linear spaces are linked the very important linear applications between linear spaces, and we see many new concepts linked to them like: kernel, image,… a linear map has many good proprieties, like continuity, that we will use all the time.

Chapter 4

When the dimension of our linear space is finite, then every linear map can be presented with a matrix. Also a matrix is seen as another facet of a linear function. Here we encounter eigenvalues and eigenfunction of matrices, the conjugate of a matrix , symmetric and **Hermirian** matrices, …. Linear functions and their associated matrices are special cases of linear operators.

Contents

1) Linear spaces: ……………………………………………………………………………………..………8

2) Norms in linear spaces :……………………..…………………………………………………22

3) Linear maps: ………………….………………………………………………………..…………37

4) Linear maps as matrices: ……………………………………………………………………..45

Chapter 1

Linear spaces

Linearity is a useful and appropriate propriety in many occasions here, and when we work on a problem involving a space of objects (numbers, functions,…)we wish that our environment ,i.e. the space in question, is linear because the things are much more simpler. As an example of how linearity is very natural in everyday life, suppose you buy 2 kilograms of apples, knowing the price of $1kg$, you only have to double the price, following the law (the linear map) for this

$$f(x+x)=f(x)+f(x)$$

or simply

$$f(2x)=2x$$

which is the law of ***homogeneity***. Imagine now if the prize law is not linear!!!!! An example of nonlinear law we propose

$$g(x) = x^2$$

So if you square this function in $x + x$ or $2x$ you will get

$$x^2 + 2x + 1$$

The worse is that the residue term $x^2 + 1$ alone is much bigger

$$x^2 + 1 > 2x$$

try now the two laws : $log(x)$ and e^x.

Vectors:

The natural origin of vectors is the plane $\mathbb{R}^2$ and also the physical space $\mathbb{R}^3$. In the general space $\mathbb{R}^n$, a vector v is determined by its components or coordinates, so in $\mathbb{R}^2$ and $\mathbb{R}^3$, we can write either

$$v = \binom{a_1}{a_2} \qquad \text{or} \qquad v = (a_1, a_2, a_3)$$

where $a_1, \dots, a_n$ are the coordinates of the vector v. By this there is no objection when we generalize vectors by coordinates to $\mathbb{R}^n$, so we likewise write

$$v = \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \qquad \text{or} \qquad v = (a_1, \dots, a_n)$$

The vector v can be seen as a *point* in the space $\mathbb{R}^n$. There is no difference between the two above notations. The only difference rises when we want to make a product of two vectors or a product of a vector and a matrix.

Orthogonality will be needed in the subsequent chapters, so we say that two vectors are *orthogonal*(or perpendicular) if their *dot product* (or scalar product) equals to zero

$$v_1 \circ v_2 = 0$$

And this presented in terms of each vectors components by the known product

$$(a_1, \dots, a_n) \circ \begin{pmatrix} b_1 \\ . \\ . \\ . \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i \times b_i = 0$$

Linear space

A *space* in mathematics is a non-empty set with a structure defined in it for certain purpose. Geometers worked in terms of linear spaces(or vector spaces) since Antiquity; but these spaces began to see their systemization as an achieved theory only since the beginning of the XX century. Very important spaces that we will need in subsequent treatment are naturally linear spaces: **Lebesgue** spaces $L^p(\Omega)$, the space of test functions $\mathcal{D}(\Omega)$, the space of distribution $\mathcal{D}'(\Omega)$ … without linearity our work would be impossible.

Definition

Let $\mathbb{K}$ a set of numbers λ_i , that is a commutative field $\mathbb{K}$ and E a set of vectors that verify two sets of axioms:

E itself is a commutative group:

i) $\forall x, y \in E : x + y = y + x = z \in E$.

ii) $(x + y) + z = x + (y + z)$.

iii) $\exists 0 \in E : x + 0 = 0 + x = x$.

iv) $\forall x \in E, \exists x' : x + x' = x' + x = 0$.

between E and $\mathbb{K}$

i) $\lambda_1.(\lambda_2.x) = \lambda_1.\lambda_2.(x)$

ii) $\lambda.(x + y) = \lambda.x + \lambda.y$.

iii) $(\lambda_1 + \lambda_2).x = \lambda_1.x + \lambda_2.y$.

iv) $1x = x$

So in a linear space we have :

i) a vector addition " + " between the vectors.

ii) a scalar multiplication " . " between the vectors and the numbers(scalars) of the field $\mathbb{K}$.

In almost the cases $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. We say that $(E, +)$ is a linear space over the field $(\mathbb{K}, +, .)$. Any field $(\mathbb{K}, +)$ is a vector space over itself $(\mathbb{K}, +, .)$.

The above axioms were written down for the first time by Polish Stephan **Banach**(1892-1945) in 1932 and shortly after by Norbert **Wiener**(1894-1964). We can read:

> ***Fréchet*** *was very excited by the fact that* ***Banach*** *had given a system of axioms for vector spaces months before* ***Wiener***"[1].

Examples:

1) The excellent linear space over itself is $\mathbb{R}$.

2) The plane $\mathbb{R}^2$ the set of vectors of geometry, makes itself a linear space.

---

[1] *Oeuvres de* ***Banach*** , prefaced by his friend H. **Steinhaus**(p.15). See *Geometry by its history* (p.262).

3) $\mathbb{R}^3$ a very special set of vectors used in physics and geometry, it is a linear space, and the appellation *vector* comes from this space.

4) in fact all $\mathbb{R}^n$ when $n > 3$ is finite are linear spaces. So $\mathbb{R}^n$ is seen as the set of all n-tuples

$$(x_1, \dots, x_n)$$

5) The set of sequences $u \coloneqq \{u_n\}$ noted ℓ^p where $1 \leq p < +\infty$, such that

$$\sum_{n=1}^{+\infty} |u_n|^p$$

converges is a linear space. In particular ℓ^∞ is the linear space of bounded sequences.

6) The set of sequences $\{u_n\}$ converging to 0, noted c_0 and the set of sequences $\{u_n\}$ that are null for a certain $n \geq k$ noted c_{00} are both linear spaces.

Remarks

− One should make a difference between the number $0 \in \mathbb{K}$ and the vector $\mathbf{0} \in E$[2].

Linear subspace

Inside any linear space E we may find another linear space, this new space is called a linear *subspace*, we can consider E as a subspace of itself.

---

[2] Some authors write the vector $\mathbf{0}$ as $\vec{\mathbf{0}}$.

Definition

Let E a linear space and F is subset of E. We say that F is a linear subspace of E if:

$i)$ For every $v_1, v_2 \in F$ we have : $v_1 + v_2 \in F$.
$ii)$ For every $v \in F$ and every $\alpha \in \mathbb{K}$ we have : $\alpha . v \in F$.

A subspace of a linear space cannot be empty, it contains at least the unit vector $\mathbf{0}$ for the operation " $+$ ". So in any linear space we can find at least two linear subspaces.

Examples:

1) a line passing by $(0,0)$ is a subspace of the plane $\mathbb{R}^2$.

2) the line $\mathbb{R}$ is a subspace of the plane $\mathbb{R}^2$ which is itself is a subspace of $\mathbb{R}^3$.

3) The set of continuous functions over $[0,1]$ noted by the notation $C^0([0,1])$ is a linear subspace of **Lebesgue** linear space $L^p([0,1])$ which itself is a linear subspace of the more general linear space of arbitrary functions $\mathcal{F}([0,1])$ so we may write

$$C^0([0,1]) \subset L^p([0,1]) \subsetneq \mathcal{F}([0,1])$$

Basis

Before explaining what a basis of a linear space is, we first define the following:

Definition
We say that n vectors in a linear space E

$$v_1, \dots, v_n$$

are *linearly independent* if the zero equation

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

implies that all the coefficients $\alpha_i = 0$.

This is extremely essential in linear algebra, linearly independent means that we cannot express any vector in terms of the other vectors (independence). From the equation

$$2v_1 - 5v_2 = 0$$

we infer that v_1 and v_1 are linearly dependent as we can write the vector v_1 in term of the other v_2 (and verse versa)

$$v_1 = \frac{5}{2} v_2$$

Now let

$$e_1, \dots, e_n$$

a set of n linearly independent vectors in E

Definition

We say that $(e_1, \dots, e_n)$ is a basis for the linear space E if every vector v in E can be written as a *linear combination* of the vectors $(e_1, \dots, e_n)$

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

We say that $(e_1, \ldots, e_n)$ *spans* the linear space , or that $(e_1, \ldots, e_n)$ is a *span* for E.

We have the convention

$$\text{span}\,(\mathbf{0}) = \{0\} \qquad , \quad \text{span}(\emptyset) = \{0\}$$

Definition:

So a set of vectors $(e_1, \ldots, e_n)$ constitute a basis for a linear space E iff:

$*)$ $(e_1, \ldots, e_n)$ spans E.

$*)$ $(e_1, \ldots, e_n)$ is linearly independent.

a basis help us to write these vectors as the alphabet letters help us to write words. The idea of a basis comes from the vector space $\mathbb{R}^n$ where any point $p \in \mathbb{R}^n$ has the form

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

where α_i are real numbers.

Remarks

$-$The basis of a given linear space is not unique, but for any vector v written in a given basis, the numbers $\alpha_1, \ldots, \alpha_n$ are in $\mathbb{K}$ and are *unique*.

– To show that a set of n vectors $(e_1, \dots, e_n)$ is a basis for $\mathbb{R}^n$, it suffices to show that the determinant

$$|(e_1, \dots, e_n)| \neq 0$$

Examples:

For the real spaces $\mathbb{R}^{1,2,3}$ we can use the following basis

i) for $\mathbb{R}$ the basis

$$(0,1)$$

ii) for $\mathbb{R}^2$

$$\{(0,1), (1,0)\}$$

iii) for $\mathbb{R}^3$

$$\{(1,0,0), (0,1,0), (0,0,1)\}$$

For the general $\mathbb{R}^n$

$$\{(1,0, \dots ,0), (0,1, \dots ,0), \dots , (0,0, \dots ,1)\}$$

The previous chosen basis are called *canonical* or *standard* as they are the usually natural used basis for these spaces

Theorem

Every linear space has a basis.

We can prove this result and many others by using the *axiom of choice*. This axiom was first stated by German mathematician **Zermelo**(1871-1953) in 1904, and although it is extremely important, there was time when it was very controversial because it implies *the existence* of a certain *function of choice f* of which mathematicians were not unanimous. Now the game is over, the axiom is accepted by all mathematicians and used everywhere in Analysis because of the following:

$\ast$) It is quite natural, and in fact, it has been used by many mathematicians from the XIX century before **Zermelo** without realizing it.

$\ast$) It is necessary for the proof of important coming results in this course like the **Hahn-Banach** theorem…

$\ast$) Since its introduction and use, no one has found a contradiction to it.

$\ast$) Logicians results have given credence to it, beginning by **Gödel** result(1938) that the axiom of choice is consistent, in a sense that it cannot be a contradiction in set theory, then by **Cohen**(1963) that the axiom of choice is independent, meaning it is not a theorem for it either[3].

$\ast$) It is equivalent to many basic assertions in other domains of mathematics as: ***Hausdaurf*** *maximum principle*, *well-ordering principle* and ***Zorn*** *lemma*. On **Zorn** lemma, H. **Brezis** says:

> *"**Zorn**'s lemma follows from the axiom of choice, …[it] has many important applications in analysis. It is a basis tool in proving some seemingly innocent existence statements such as "every vector space has a basis" and " on any vector space there are nontrivial linear functional". Most analysts do not know how to prove **Zorn**'s lemma; but it is quite essential for an analyst to understand the statement of*

---

[3] The interested reader in the history of this axiom can look in the book of G.H. **Moore**.

Zorn's lemma and to be able to use it properly!"[4]

Dimension of a linear space

The idea of the *dimension* of a vector space E comes from the usual $\mathbb{R}^n$. It is linked to:

i) the concept of *linear independence* .

ii) the number of linearly independent vectors in E.

In fact, the dimension of a linear space is expressed as the *maximum*[5] number of vectors that its basis contains.

Definition

let E be a linear space, the dimension of E is the maximum number of independent vectors in E. If this number is n we write

$$\dim(E) = n$$

So we have for any linear space E with a basis $(e_1, \dots, e_n)$:

i) We cannot write any vector from the space E in less than n linearly independent vectors.

ii) If we write a vector v in a linear space of dimension n in $n+1$ vectors, these $n+1$ vectors must be linearly dependent.

---

[4] In his book : *Functional analysis* p. 2.

[5] So if there are no maximum bound for the number of vectors the dimension is infinite.

Examples:

1) The line $\mathbb{R}$ for example has dimension 1 because any vector has one component

$$v = a_1 e_1$$

Where any $e_1 \neq 0$ is linearly independent because if $a_1 e_1 = 0$ then $a_1 = 0$.

2) The plane $\mathbb{R}^2$ has dimension 2 because any vector has two components and

$$v = a_1 e_1 + a_2 e_2$$

where (e_1, e_2) are two chosen linearly independent vectors.

Theorem

let E be a linear space, and V a subspace of E

$$\dim(E) \geq \dim(V)$$

So

$$\dim(\mathbb{R}^3) > \dim(\mathbb{R}^2) > \dim(\mathbb{R})$$

Defining the dimension of a linear space, then the linear spaces with which we deal can be of two categories:

−*Finite dimension spaces:* which are excellently presented by the original $\mathbb{R}^n$, the dimension n is the maximum number of linearly independent vectors.

−*Infinite dimensional spaces*: mathematicians at the end of the XIX century ,have discovered that when dealing with differential equations, one can consider linear spaces with an infinite basis, and these new spaces proved to be very tricky. These are examples of linear spaces with infinite dimension, they are spaces of functions that verify certain conditions:

i) $C[a,b]$: space of continuous functions over the interval $[a,b]$.

ii) $L^p[a,b]$ **Lebesgue** spaces of functions defined over $[a,b]$.

iii) $W^{p,n}[a,b]$: **Sobolev** spaces of functions defined over $[a,b]$.

Exercises

1) Show that two vectors in the plane $\mathbb{R}^2$ are dependent iff they are parallel.
1) show that $\mathbf{0}$ is a vector space, it is a subspace of any linear space.
1) Show that a linear subspace cannot not empty and contain at least the unit element **0**.
1) Show that the vector $\mathbf{0}$ is dependent.
1) Show that a set of vectors containing $\mathbf{0}$ are dependent.
1) What are the components of $\mathbf{0}$?
1) Show that a set of vectors which contains a linearly dependent subset are linearly dependent.
1) are the following linear subspace of $\mathbb{R}$?
i) the interval $]0,+\infty]$.
ii) the interval $[0,+\infty]$.
iii) the interval $[0,1]$.
1) show that $\{\mathbf{0}\}$ is a subspace.
1) show that the vector $\mathbf{0}$ belongs to any subspace.
1) is the line given by the equation $y = x + 1$ a linear subspace of $\mathbb{R}^2$?

1) Show that the components of a vector in a given basis are unique.

1) what is the minimum numbers of linearly independent vectors in the linear spaces $\mathbb{R}, \mathbb{R}^2$ and $\mathbb{R}^3$?

1) give examples of each sequence of the following sets:

i) ℓ^1

ii) ℓ^2

iii) c_0

iv) c_{00}

1) show that : $c_{00} \subset \ell^p \subset c_0 \subset \ell^\infty$.

1) is the following vectors in $\mathbb{R}$ form a basis

i) $((1,0))$

ii) $((2,3))$

iii) $((0,0))$

1) is the following vectors in $\mathbb{R}^2$ form a basis

i) $((1,0), (1,0))$

ii) $((1,0), (1,3))$

iii) $((1,2), (-1,1))$

1) is the following vectors in $\mathbb{R}^3$ form a basis

i) $((1,0), (1,-1)(1,1))$

ii) $((-1,0), (1,0), (2,-2))$

iii) $((1,1), (2,1), (-1,1))$

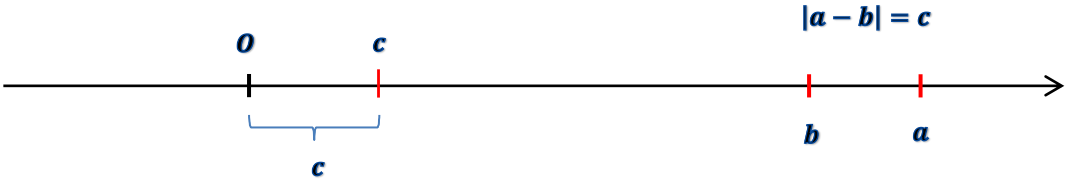
Chapter 2

Norms

When we deal with real numbers in the real line $\mathbb{R}$ to make a comparison between any two numbers we take their distance to each other. We can present the distance between two numbers a and b by a *positive number* given by the absolute value

$$|a-b| = |b-a| \geq 0$$

and this positive number is in fact the distance from 0, so the difference between the numbers can be moved to the origin $\boldsymbol{O}$ specified by the number 0.



In many situations that we met treating **Cauchy** problem, we have to compare between two objects in our linear spaces: two vectors, two matrices, two functions …as we compare between lengths, then mathematicians had to find tools that permit them to do such operations, and this perfect tool is the concept of *norm* , the generalization of usual *length* in $\mathbb{R}^{1,2,3}$.

Topology

This is just a short review of general topology , and we insist that topology is all about continuity.

Definition

Let τ a set containing subsets of E that verifies:

$i)$ E and $\emptyset \in \tau$.
$ii)$ $\bigcup_{i\in I} A_i \in \tau$.
$iii)$ $\bigcap_1^n A_i \in \tau$.

We say that τ is a topology over E, a topology is defined by neighborhoods called "open sets"(as open intervals are neighborhood for the real numbers). A general neighborhood is a very abstract form of the usual distance that we know, it translates the expression "a is near to b" without reference to any proper distance ,length,… so to get richer neighborhoods we move to the more concrete metrics.

Metric

A metric defined in a set, is stronger than a topology:

Definition

We say that d is a metric over a set E if:

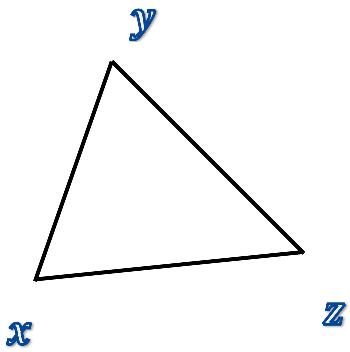
$i)$ $d(x,y) \geq 0.$
$ii)$ $d(x,y) = 0 \Leftrightarrow x = y.$
$iii)$ $d(x,y) = d(x,y).$
$iv)$ $d(x,z) \leq d(x,y) + d(y,z).$

To express a distance on a set , the two following proprieties are sufficient[6]

$*)$ $d(x,y) = 0 \Leftrightarrow x = y.$

$*)$ $d(x,z) \leq d(x,y) + d(y,z).$

But in common references of Functional Analysis we find two more($i)$ and iii). The propriety (i) is easy to prove, for the propriety (iii) we can prove it like this : we start from



$$d(y,z) \leq d(x,y) + d(x,z).$$

If $x = z$ then

$$d(y,z) \leq d(z,y)$$

then if we start with

$$d(z,x) \leq d(x,y) + d(y,z).$$

and putting $x = y$ then

$$d(z,y) \leq d(y,z)$$

---

[6] See **Banach** (p. 8) referring to A. **Lindenbaum**. In fact **Lindenbaum** (1904-1941) said in his paper:" There are still two propositions: $\rho(x,y) \geq 0$ and $\rho(x,y) = \rho(y,x)$, which we see separately postulated elsewhere, this because the triangular inequality is almost always expressed in a less suitable form… ".

The result follows■.

Every set E with a metric defined on it is a *metric space* , and is noted (E,d) so

$$metric\ spaces \ \subsetneq\ topological\ spaces$$

Any metric *induces* a topology, this means there is a topology or a system of neighborhoods associated to the metric as a metric is a special case of a topology.

Examples:

1) We return to the metric

$$d(x,y) = |x-y|$$

defined on the real line $\mathbb{R}$. The conditions of a metric are fulfilled: the first is by definition positive $|x-y| \geq 0$, and we have also $|x-y| = 0$ iff $x = y$. For the third we have

$$|x-y| = |-1||y-x| = |y-x|$$

For the triangular inequality we have

$$|x-z| = |(x-y)+(y-z)| \leq |x-y| + |y-z|$$

2) The second example is very special: it is the natural distance in the plane $\mathbb{R}^2$, it is measured by the famous **Pythagoras** theorem. Let $x(a_1,b_1)$ and $y(a_2,b_2)$ two points(vectors) in $\mathbb{R}^2$, then we define

$$d(x,y) = \sqrt{(a_2 - a_1)^2 + (b_2 - b_1)^2}$$

When we generalize to the $\mathbb{R}^n$ is

$$d_p(x,y) = \sqrt[p]{\sum_{i=1}^{n} (a_i - b_i)^p} = \left(\sum_{i=1}^{n} (a_i - b_i)^p \right)^{1/p}$$

It is the *canonical* or *usual* metric in $\mathbb{R}^n$. The reader may have noticed that the first example is just a special case of this, for we have

$$d(x,y) = \sqrt{\sum_{i=1}^{1} (a_i - b_i)^2} == \sqrt{(a-b)^2} = |a-b| = |x-y|$$

3) We now give freedom to our imagination and consider a distance called the *discrete metric*, its defined as

$$d(x,y) = \begin{cases} 1 & if \quad x \neq y \\ 0 & if \quad x = y \end{cases}$$

We see immediately that) , $ii)$ and $iii)$ are verified. We look to prove the triangular inequality

$$d(x,z) \leq d(x,y) + d(y,z)$$

$-$ if $d(x,z) = 1$ (meaning that $x \neq z$) then $d(x,y)$ and $d(y,z)$ cannot be both 0, because that leads to a the contradiction $x = z$.

$-$ if $d(x,z)=0$, then whatever the value at the right is, the inequality is true ■.

For this bizarre *discrete metric*, all the elements in E are equidistance.

Norms

Every norm is a special metric, norms are the special metrics that we will work with all the time.

Definition

We say that $\| \quad \|$ is a norm if:

i) $\|x\|=0 \Longrightarrow x=\mathbf{0}$.
ii) $\forall\lambda\in\mathbb{K}: \ \|\lambda x\|=|\lambda|\|x\|$
iii) $\|x+y\|\leq\|x\|+\|y\|$

The first axiom is the *definiteness* of the norm and it is usually written

$$\|x\|=0 \Leftrightarrow x=\mathbf{0}$$

the opposite implication is provable. The second axiom is *homogeneity* of the norm. The third is of course is the known triangular inequality.

Every set E with a norm $\| \quad \|$ defined on it is a *normed space*, noted $(E,\| \quad \|)$ and we have

$$normed\ spaces \ \subsetneq \ metric\ spaces$$

So every norm is a metric and every metric is a topology, both metrics and norms define abstracts distances. Like metrics, for any norm there is a topology (a system of neighborhoods) associated to it

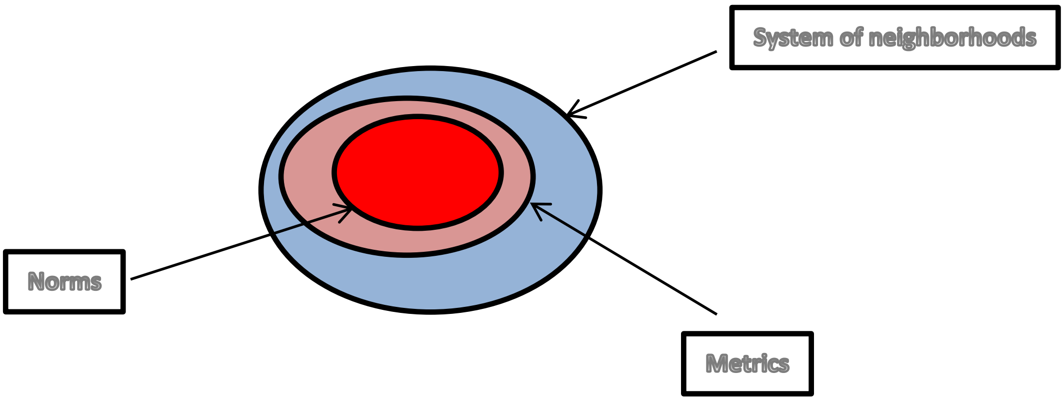


Fig. Unifying the three notions.

A norm $\| \quad \|$ defines a stronger topology and helps us to define and work with the notion of *convergence*. It is hard usually to find the right norm (especially if the space is infinite dimensional), we will discover that it belongs to the genius of a mathematician when working on a new space of functions to choose the right norm in order that the space is complete...

As we have defined the norm we can calculate *the distance* between two vectors(or objects) by using the metric

Definition:

The distance between two objects(vectors,…) x and y is

$$d(x,y) = \|x - y\|$$

We can also treat convergence by the norm, so a sequence of maps f_n converge to a certain function f in terms of the norm

$$\|f - f_n\| \longrightarrow 0$$

and as

$$d(x,0) = \|x - 0\| = \|x\|$$

norms are "weights" of objects and the distance is like the difference of those weights. Particularly the norm of x is the distance of x from the object 0(the point 0, the vector 0, the function 0,…)

Examples:

1) We return the usual metric

$$\|x\| = |x|$$

i) $\|0\| = |0| = 0$, and for ii) $\|x\| = |\lambda x| = |\lambda||x|$. iii) We simply use the triangular inequality in the plane. The distance between $x = -5$ and $y = 5$ on the real line $\mathbb{R}$ is

$$d(x,y) = \|5 - (-5)\| = \|-5 - 5\| = 10$$

2) Comparatively to the already seen metrics we can define the corresponding general standard norms

$$\| \quad \|_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$$

and the usual **Euclidian** for $p = 1,2$.

3) We can also borrow the *supremum norm* form the uniform convergence of functions to numbers, so we define the norm

$$\|u\|_\infty = max\{|u_n|\}$$

on the set of sequences, and sup becomes max as the u_n are defined numbers[7].

Equivalence of norms

With all the previous norms there is a question that asks itself: these norms all work, so are there some differences between them?

So we say that two norms $\|u\|_1$ and $\|u\|_2$ are *equivalent* if there exist two positive numbers α and β such that

$$\beta\|u\|_2 \leq \|u\|_1 \leq \alpha\|u\|_2$$

as

Theorem

In finite dimension spaces, all norms $\| \quad \|$ are equivalent.

---

[7] a function may not reach its max so we should put sup in place of max.

This result is false in infinite dimensional spaces.

Semi-Norms

In many occasions we met metrics such that the first propriety of a norm cannot be verified, namely:

$$\|x\| = 0 \Longrightarrow x = \mathbf{0}.$$

So in this case we are in face of a semi-norm.

Definition

We say that $\| \quad \|$ is a semi- norm if:

i) $\forall \lambda \in \mathbb{K} : \ \|\lambda x\| = |\lambda| \|x\|$
ii) $\|x + y\| \leq \|x\| + \|y\|$

Example:

As an example of a semi-norm we consider points $z = (x, y)$ in the plane $\mathbb{R}^2$ and we define the norm

$$\|z\| = \|(x, y)\| = |x - y|$$

we see that $|x - y| = 0$ whenever $x = y$.

As a very known example in Analysis, we can cite **Lebesgue** integrable function space $\mathcal{L}(\Omega)$ (not $L(\Omega)$), which constitutes a semi-normed space. We will see also with the spaces $L^p(\Omega)$ the concept of *quasi-norms*.

Remark

– Using the unifying language of mathematics, a metrics is seen as application a from the linear space E to the positive real numbers

$$d(a,b) : E \times E \longrightarrow \mathbb{R}^{+}$$

Likewise for the norms and semi-norms

$$\|a\| : E \longrightarrow \mathbb{R}^{+}$$

Famous inequalities

As J. **Dieudonné** says :

> *"Infinitesimal Calculus is the learning of the manipulation of inequalities much more than equalities".*

In Functional Analysis, mathematicians have to compare between his objects of study like: vectors, matrices, functions, …So after we have found the "weight" which is the concept of *norm*, so we now need also to know the "tools"(the scales) to implement this weight, these tools are the *inequalities* . What is very important about inequalities is that they are everywhere in mathematics not only in Analysis, and in this broad domain we find many interesting relations[8]. The big challenge that remains for the reader is : when and how to use an inequality.

---

[8] English mathematician **Hardy**(1877-1947) et al. have written a whole book on inequalities in more than 300 pages. The seminal proof of the ***Poincaré** conjecture*, due to Gregory **Perelman**, extensively uses ***Harnack** inequality*.

Minkowski inequality

Due to German Hermann **Minkowski**(1864-1909). For real numbers a_i, b_i we have

$$\left(\sum_{i=1}^{n}|a_i+b_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n}|a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n}|b_i|^p\right)^{1/p}$$

Minkowski inequality uses only one number p and it implies an addition. A special case is the triangular equality, it is the case when $p = n = 1$

$$|a+b| \leq |a| + |b|$$

Also in spaces of functions we can write it

$$\left(\int_a^b |f(x)+g(x)|^p\,dx\right)^{1/p} \leq \left(\int_a^b |f(x)|^p\,dx\right)^{1/p} + \left(\int_a^b |g(x)|^p\,dx\right)^{1/p}$$

Hölder inequality

Due to German Otto **Hölder** (1859-1937). Unlike **Minkowski** inequality, this is associated with two numbers (p, q) which are *conjugated*, namely they verify

$$\frac{1}{p}+\frac{1}{q}=1$$

So we have the **Hölder** inequality

$$\sum_{i=1}^{n}|a_i.b_i| \le \left(\sum_{i=1}^{n}|a_i|^p\right)^{\frac{1}{p}} . \left(\sum_{i=1}^{n}|b_i|^q\right)^{\frac{1}{q}}$$

notice that it is a product. Special case of this inequality is the very known **Cauchy-Schwartz** inequality that holds when $p = q = 2$:

$$\sum_{i=1}^{n}|a_i.b_i| \le \left(\sum_{i=1}^{n}|a_i|^2\right)^{\frac{1}{2}} . \left(\sum_{i=1}^{n}|b_i|^2\right)^{\frac{1}{2}}$$

This special equality is also called ***Bouniakowsky** inequality*, to Ukrainian Viktor **Bouniakowsky**(1804-1889) who discovered it in 1859.

We also have the integral version for functions f integrable on $[a, b]$

$$\int_a^b |f(x)g(x)|dx \le \left(\int_a^b |f|^p\,dx\right)^{\frac{1}{p}} . \left(\int_a^b |g|^q\,dx\right)^{1/q}$$

Both **Minkowski** and **Hölder** inequalities are called *convexity inequalities*.

Young inequality

Due to English mathematician W. **Young**(1863-1842), with conjugated (p, q) as usual

$$\frac{1}{p} + \frac{1}{q} = 1$$

And for every $a, b \ge 0$ we have

$$\frac{a^p}{p}+\frac{b^q}{q}\geq ab$$

Exercises

1) is $\{\mathbb{R}, \emptyset, all\ \ ]a,b[\}$ a topology over $\mathbb{R}$?

1) show that any metric verifies $d(x,y)\geq 0$.

1) say if the following are metrics**:**

i)

$$d(x,y)=\begin{cases}1 & if\ \ x=y\\ 0 & if\ \ x\neq y\end{cases}$$

1) show that any metric given by a norm should verify:

$i)\ d(\alpha x,\alpha y)=|\alpha|d(x,y)$.

$ii)\ d(x+\alpha,y+\alpha)=d(x,y)$.

1) show that any norm verify the axioms of a metric.

1) let $\|\ \ \ \|$ be a norm, show that : $x=\mathbf{0}\Longrightarrow\|x\|=0$.

1) is the discrete metric a norm?

1) prove the following inequalities :

i) $\|y-x\|=\|x\|+\|y\|$

i) $\|x\|-\|y\|=\|x-y\|$

1) Say if the following are only semi-norms on $\mathbb{R}^2$:

i) $\|(x,y)\|=|x+y|$

i) $\|(x,y)\|=|x|$

i) $\|(x,y)\|=|\alpha y|$, $\alpha\in\mathbb{R}$.

1) write the expression of the standard norm $\|\ \ \ \|_p$ for $p=1$ and $p=2$.

1) show that **Minkowski** inequality is true for $p=1$.

1) show that

$$d(x,y)=\sqrt{(a_1-b_1)^2+(a_2-b_2)^2}$$

is a metric in $\mathbb{R}^2$.

1) Show that

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

is a metric in $\mathbb{R}^n$.

1) show that the following is a norm

$$\|x\|_2 = \sqrt{(a_1)^2 + (a_2)^2}$$

Chapter 3

Maps between linear spaces

We now introduce maps between two linear spaces, and many concepts associated to them. The first thing to know is that : if f is a linear map from E to F ,then both E and F must be linear spaces.

Linear map

A linear map[9] from E to F transforms a vector from E to another vector in F , the name linear map is an appellation to any *linear application* , *linear function* or a linear transformation .

---

[9] We say also: mapping, application, function.

Definition

We say that $f: E \to F$ is a *linear map* if :

$i)$ $f(x+y) = f(x) + f(y)$

$ii)$ $f(\alpha x) = \alpha f(x)$, $\forall \alpha \in \mathbb{K}$

We can simply compose the two conditions $\forall\, x, y \in E$ as

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

when the scalars α, β are in $\mathbb{K}$ as $\mathbb{R}$ or $\mathbb{C}$.

Examples:

1) Any linear map from $\mathbb{R}$ to $\mathbb{R}$ has the form[10]

$$f(x) = ax$$

where is a a real number.

2) The map from $\mathbb{R}^2$ to $\mathbb{R}$

$$f\begin{pmatrix} x \\ y \end{pmatrix} = (x + 2y)$$

3) The map from $\mathbb{R}$ to $\mathbb{R}^2$

$$f(x) = \begin{pmatrix} 2x \\ 3x \end{pmatrix}$$

---

[10] In some references aimed at engineering and applied mathematics, and with a non-rigorous fashion that makes particularly purists cry, one can find expressions like :" $y = ax + b$ is a linear function", meaning an equation for a line.

4) The map from $\mathbb{R}^2$ to $\mathbb{R}^2$

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix}$$

Remark

$*$) linear maps are a special case of *morphisms* (or *homomorphisms*) between groups (as $(E,+)$ and $(F,+)$ are groups) that are defined as

$$f:(G_1,\circledast) \longrightarrow (G_2,\odot)$$
$$f(v_1 \circledast v_2) = f(v_1) \odot f(v_2)$$

$*$) A linear map is very desirable map which has many special proprieties like:

$$f(0) = 0$$

A law that follows from a propriety of any morphism of groups

$$f(e_1) = e_2$$

$*$) When f is a morphism from E in itself, it is called an ***endomorphism***. When the morphism f is bijective, it is called an ***isomorphism***. If f is an isomorphism from E to itself, then f is called an ***automorphism***.

The two examples 1) and 4) are endomorphisms.

Kernel and images

To a linear maps between two linear spaces E and F is associated two subspaces: the *null subspace* (or *kernel*) and the *image subspace*.

Definition

Let $f: E \rightarrow F$ a linear map then we can define two sets

$i)$ The elements $x \in E$ where $f(x) = 0$ is called ***kernel*** of f, it is a linear subspace of E.

$ii)$ The elements $y \in F$ such that $f(x) = y$ is the ***image*** space , and it is a linear subspace of F.

We can prove this propriety about the kernel:

Theorem:

if $Ker f = \{0\}$ if and only if then f is injective .

Proof: :

$i)$ $Ker f = \{0\} \Longrightarrow f$ is injective.

We say that f is injective when

$$f(x_1) = f(x_2) \quad \Longrightarrow \quad x_1 = x_2$$

Suppose that $Ker f = \{0\}$ only, then if $f(x_1) = f(x_2)$ then $f(x_1 - x_2) = 0$ (as f is linear) which gives automatically $x_1 = x_2$ as 0 is the only value that makes f vanishes, so f is injective ■.

When the dimension of the spaces E is finite these two subspaces has interesting proprieties

Theorem:

let $f: E \longrightarrow F$ a linear map and E is a linear space of finite dimension then

$$dim(E) = dim(Ker) + dim(Im)$$

Examples:

1) consider the endomorphism

$$f\binom{x}{y} = (0, y)$$

The kernel $(x, 0)$ is a whole line and its dimension is 1, the image $(0, y)$ is a whole line and its dimension is 1, and we have

$$dim(E) = dim(Ker) + dim(Im) = 1 + 1 = 2$$

Dual space

Definition

We call a *dual* two spaces:

$i)$ The algebraic dual of E: it is the set the of linear maps defined on E , it is noted : E^*

ii) The topological dual of E: the set of continuous maps defined on E, it is noted E'.

The theory of the duals of linear spaces was developed in the beginning of the XX century, and it has been used for example by Laurent **Schwartz** (1915-2002) to build his theory of distributions.

Continuity of linear maps

If the set E is a finite dimensional linear space we have

$$f \in E^* \implies f \in E'$$

That means:

Theorem

In a finite dimension space E, all linear maps are continuous.

In a space with an infinite dimension, this may be not true; there are maps defined on an infinite dimensional space, although linear, but are not continuous, as we will see in the theory of linear operators.

Exercises

1) say if the following maps are linear:

$i)$ from $\mathbb{R}^2$ to $\mathbb{R}$

$*)\ f\begin{pmatrix} x \\ y \end{pmatrix} = (x - y + 1).$

$*)\ f\begin{pmatrix} x \\ y \end{pmatrix} = (x - xy)$

$*)\ f\begin{pmatrix} x \\ y \end{pmatrix} = (2x - \sqrt{3}y)$

$ii)$ from $\mathbb{R}$ to $\mathbb{R}^2$

$*)\ f(y) = \begin{pmatrix} 1 - y \\ y \end{pmatrix}$

$*)\ f(y) = \begin{pmatrix} log(y) \\ y + \sqrt{2} \end{pmatrix}$

$*)\ f(y) = \begin{pmatrix} \pi - y \\ \frac{1}{y} \end{pmatrix}$

$iii)$ From $\mathbb{R}^2$ to $\mathbb{R}^2$

$*)\ f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y + 1 \\ e^{x+y} \end{pmatrix}$

$*)\ f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 6y \\ xe^{y} \end{pmatrix}$

$*)\ f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{x} - y \\ \frac{x+y}{log(2)} \end{pmatrix}$

1) Show that if f is a linear map from E to F both E and F must be linear spaces. Is the converse true?

1) let f be a linear bijective map, show that f^{-1} is linear.

1) Show that $f(0) = 0$.

1) Show that a morphism of groups f verifies : $f(e_1) = e_2$.

1) Give a nonlinear morphisms from $(\mathbb{R}^{*+}, \times)$ to $(\mathbb{R}, +)$.

1) Give a nonlinear morphisms from $(\mathbb{R}, +)$ to $(\mathbb{R}, \times)$.

1) are the functions $ln(x)$ and e^x isomorphisms?

1) Show that if f is a linear injective then $Kerf = \{0\}$.

1) is the following map

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\x-y\end{pmatrix}$$

an isomorphism? an automorphism?

1) is the map

$$f\begin{pmatrix}x\\y\end{pmatrix}=(0,y)$$

injective? why?

Chapter 4

Linear maps as matrices

Linear maps and matrices are very linked to each other, the set of linear maps is a linear space and the set of matrices is also a linear space over $\mathbb{R}$ and $\mathbb{C}$. Moreover we will discover here that any linear map between two linear spaces can be presented as a matrix, which is not necessary squared.

Matrix

The theory of matrices has been introduced by **Cayley**(1821-1895), the word *matrix*[11] was coined by his friend **Sylvester**(1814-1897) , one of the purpose of this theory also was to facilitate the writing of linear transformations(or linear systems)

---

[11] A Latin word meaning *womb*.

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \cdot \\ b_n \end{pmatrix}$$

by matrix notation we can write it as

$$A \begin{pmatrix} x_1 \\ \cdot \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \cdot \\ b_n \end{pmatrix}$$

Or simply

$$Ax = b$$

where a_{ij} are the coefficients, the $x = (x_1, \dots, x_p)$ are the unknowns and $b = (b_1, \dots, b_n)$ the second member. So to study a big cumbersome linear system of equations, one has only to study the matrix A.

Remark

The existence of solutions for any linear system $Ax = b$ depends on the determinant of the squared matrix A:

$*)$ if $det(A) \neq 0$, we have a unique solution.

$**)$ if $det(A) = 0$, we have two cases :

$-$either there are no solutions at all.

$-$or there are infinitely many solutions.

Some algebra of matrices

There is a lot to say about this subject, but we keep ourselves to the minimum:

$-$A matrix can have one of the three used notations

$$\begin{pmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{pmatrix} \qquad \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} \qquad \begin{Vmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{Vmatrix}$$

or

$$(a_{ij}) \qquad [a_{ij}] \qquad \|a_{ij}\|$$

—A matrix with n rows and m columns may be written as $n \times m$ matrix. Any vector either orthogonal or horizontal

$$\begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \qquad , \qquad (a_1, \dots, a_n)$$

is seen as a matrix of the form $n \times 1$ or $1 \times n$. A matrix A whose entries all are zeros is called the *null* or *zero matrix*.

—If the matrix A is squared, we call it simply matrix of *order n*. Any matrix of order n then , unlike arbitrary matrices, has a diagonal composed of the diagonal numbers noted a_{ii}.

—The *transpose* of a matrix A noted A^T is the matrix obtained when we invert the rows and column, so (a_{ij}) become (a_{ji}). The transpose of the vector(matrix)

$$A = (a_1, \dots, a_n)$$

is

$$A^T = \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \qquad ,$$

we have the propriety

$$(AB)^T = B^T A^T$$

$-$Any matrix of order n, possess a determinant noted $det(A)$ which may be equal to 0, we have

$$det(AB) = det(A)\,det(B)$$

$-$ If $det(A) = 0$, we say that A is *singular*, in this case A has no inverse.

$-$A product of two matrix A and B can be the null matrix without one of the two matrices is, as a general example for the 2×2 case

$$\begin{pmatrix} a & b \\ a & b \end{pmatrix} \times \begin{pmatrix} -b & -b \\ a & a \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

notice that both determinants are zero, **Kline** wrote[12]

> " ***Cayley*** *asserted that the product of two matrices may be zero without either being zero, if either one is singular. Actually,* ***Cayley*** *was wrong; both matrices must be singular. For if* $AB = 0$, $A \neq 0$, $B \neq 0$, *and only* A *is singular, then the inverse of* B, *namely* B^{-1} , *exits, and* $ABB^{-1} = 0B^{-1} = 0$. *But* $BB^{-1} = I$. *Therefore* $AI = 0$ *or*

---

[12] In **Kline** text I have put *singular* in place of *indeterminate.* See **Kline** p.807.

$$A = 0\text{ ".}$$

Matrices and linear maps:

An important result that illustrate the relation between matrix and linear maps

Theorem

To every matrix of n rows and p columns we can associate a linear map f, and vice versa.

A linear map transforms a vector to another vector, a matrix A in a linear system $Ax = b$ does the same work and with its coefficients transform a vector to another vector.

So the similarity

$$f(v) = (w) \qquad \sim \qquad A(v) = (w)$$

If the linear map f is from $\mathbb{R}^m$ to $\mathbb{R}^n$ then the matrix has the form $n \times m$ (n rows and m columns) and vice versa.

Examples:

i) The matrix has 2 rows and 2 columns

$$\begin{pmatrix} 2 & -2 \\ 1 & 3 \end{pmatrix}$$

so it is associated to a linear map from $\mathbb{R}^2$ to $\mathbb{R}^2$

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - 2y \\ x + 3y \end{pmatrix}$$

ii) The following map is from $\mathbb{R}^2$ to $\mathbb{R}^2$

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ x + y \end{pmatrix}$$

its associated matrix is $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ since

$$\begin{pmatrix} x - y \\ x + y \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

iii) the following matrix has 3 row and 1 columns

$$\begin{pmatrix} 1 \\ 3 \\ -3 \end{pmatrix}$$

its associated linear map is from $\mathbb{R}$ to $\mathbb{R}^3$ as

$$f(x) = \begin{pmatrix} 1 \\ 3 \\ -3 \end{pmatrix}(x) = \begin{pmatrix} x \\ 3x \\ -3x \end{pmatrix}$$

Remark

$*$) To a given linear map $f: E \longrightarrow F$ the associated matrix changes according to the basis in E.

Eigenvalues of a squared[13] matrix

As a matrix transforms a vector to another vector we can write the general form of this transformation as $Av = w$, among all these vectors there are special vectors associated to each matrix A of order n verifying

$$Av = \lambda v$$

the reader observes in this case that A does not change the direction of the vector v , only its scaling or its sense, so the number λ is the *eigenvalue*[14] of A and v its *eigenvector*.

There is a method to find these eigenvalues:

Definition

Let A a matrix of order n, then the solutions of the polynomial equation

$$P_n(\lambda) = det(A - I\lambda) = 0$$

are called the *eigenvalues* of A.

The polynomial $P_n(\lambda)$ is called the *characteristic* polynomial of A and its roots , the eigenvalues, are named the *specterum* of A.

Example:

---

[13] Keep in mind that a non-squared matrix has no determinant.

[14] Form the German word *Eigen* (own) so the meaning is : own-value, typical-value.

Let A the following matrix

$$\begin{pmatrix} 1 & 1 \\ 9 & 1 \end{pmatrix}$$

So we form

$$det(A-\lambda I) = det\left(\begin{pmatrix} 1 & 1 \\ 9 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

$$= \begin{vmatrix} 1-\lambda & 1 \\ 9 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 8$$

So the roots of the characteristic polynomial $\lambda^2 - 2\lambda - 8$ constitute the spectrum

$$spec(A) = \{-2\ ,4\}$$

Remark

$*$) As a linear map f and its associated matrix A are morphisms(endomorphism,...) we usually say : find the eigenvalues of the endomorphism,…

$*$) The eigenvalues of an endomorphism,… as solutions of a polynomial equation, can be either real or complex.

We have the following interesting result

Theorem

Let A be a squared matrix and let λ_i be its eigenvalues, then

$\sum_1^n \lambda_i = Tr(A) \qquad , \qquad \prod \lambda_i = det(A).$

The trace of a squared matrix A is the sum of its diagonal numbers a_{ii}

$$Tr(A) = \sum_{i=1}^{n} a_{ii}$$

In the preceding example

$$\sum_1^n \lambda_i = -2 + 4 = 2 = 1 + 1 = Tr(A)$$

and

$$\prod \lambda_i = (-2)(4) = -8 = 1 - 9 = det(A)$$

Another result linking matrix to its eigenvalues

Theorem

Any matrix A of order n verifies its characteristic polynomial which is also of degree n. So

$$\sum_{i=0}^{n} a_i A^i = 0_n$$

So for the preceded example

$$\begin{pmatrix} 1 & 1 \\ 9 & 1 \end{pmatrix}^2 - 2\begin{pmatrix} 1 & 1 \\ 9 & 1 \end{pmatrix} - 8\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The result is known as ***Cayley-Hamilton** theorem* , attributed to **Cayley**(1858) who proved it in the special case $n = 3$ and to Irish W. **Hamilton** (1805-1865) who proved it for linear transformations(1853), the general result for matrices has been proved by German **Frobinius**(1849-1917) in 1878.

Symmetric matrix

One important matrix that will be considered much is the symmetric matrix.

Definition

We say that A is a symmetric matrix if:

$$A^T = A$$

We can simplify this by writing it by the matrix coefficients : for all i, j :

$$a_{ij} = a_{ji}$$

an example of asymmetric matrix is

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$$

An interesting case of a symmetric matrix is when the coefficients a_{ji} are in $\mathbb{C}$: we say that a matrix is ***Hermitian*** if

$$\overline{A^T} = A$$

where : $\overline{a + ıb} = a - ib$, an example is

$$A = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$$

The matrix $\overline{A^T}$ has also the name *self-adjoint* of A. The reader can easily prove that every real symmetric matrix is **Hermitian**.

French mathematician C. **Hermite**(1822-1901) proved an interesting propriety for such matrices in 1855.

Theorem

If A is **Hermitian** then its eigenvalues are real.

Exercises

1) Show that : $\det(A^T A) \geq 0$.

1) Show that $(A^{-1}) = \frac{1}{\det(A)}$. What do you infer?

1) Find $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}^{-1}$ by two methods.

1) Find the determinant of the matrix $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix}$ by three methods.

1) Show that A^{-1} if it exists is unique.

1) Show that: $(AB)^{-1} = B^{-1}A^{-1}$

1) discuss the existence of solutions for the linear system : $Ax = 0$.

1) show that if λ is an eigenvalue for A , then λ^{-1} is an eigenvalue for A^{-1}.

1) if $A^2 = A$ the A is an idempotent matrix (idem : *same* potent). Show that $det(A) = 0$ or 1.

1) Show that all idempotent matrices are singular except I.

1) Let A idempotent, show that $I - A$ is idempotent. What about $A - I$?

1) if $A^{-1} = A^T$, A is orthogonal $\in O_n(\mathbb{R})$. Show that $det(A) = \pm 1$.

1) in : $A = P^{-1}BP$, A and B are similar . Show that $\det(A) = \det(B)$.

1) give the matrix associated to each linear map:

$i)$

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ x - 2y \end{pmatrix}$$

$ii)$

$$g\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z - x - y \\ x + y \end{pmatrix}$$

$iii)$

$$h\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ x + y \\ z \end{pmatrix}$$

1) give the linear maps associated to each matrix:

$i)$

$$\begin{pmatrix} 0 \\ 1 \\ -3 \\ 1 \end{pmatrix}$$

$ii)$

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$iii)$

$$(1, \quad 2, \quad -1)$$

1) Show that : $A + \ A^T$ is symmetric.

1) let A any matrix, show that AA^T is squared. Show then that AA^T is always symmetric.

1) Show that if A is symmetric then A^{-1} (if it exists) is also symmetric.

1) Show that any real symmetric matrix is **Hermitian**.